Proximity Measurement mit geometrischen Modellen

Arthur Liske, Mathematik

17. Juni 2024

In dieser Ausarbeitung geht es darum, wie man eine Metrik auf ordinalen Ähnlichkeitsdaten herleiten kann, die zusätzlich additiv ist und durch das additive difference model repräsentiert werden kann. Das Hauptresultat ist, dass die l^p Metrik die einzige additive-difference Metrik, die additiv auf Liniensegmenten ist, was eng mit der Homogenität dieser Metrik zusammenhängt. Zum Herleiten dieses Hauptresultates werden die Grundelemente von Messtheorie, sowie von proximity measurement mithilfe von geometrische Modellen eingeführt. Diese Ausarbeitung basiert auf dem Paper [6] und den Büchern [5] und [2].

Inhaltsverzeichnis

1. Einleitung

Die Grundlegende Idee von Proximity measurement ist es, jeweils zwei Stimulipaare auf deren Ähnlichkeit zu vergleichen. Zum Beispiel könnten in einem Experiment Vier Farben a, b, c, d zusammen mit der Frage "Sind sich die Farben a und b ähnlicher als die Farben c und d" Präsentiert werden. Die Aufgabe wäre dann, immer die Stimulipaare nach Ihren Ähnlichkeiten zu sortieren. Diese Ähnlichkeitsvergleiche will man anschließend durch eine Metrik repräsentieren. Dies macht man, indem man fordert, dass wenn zwei Stimuli a und b ähnlicher sind, als zwei weitere Stimuli c und d, dass dann der metrische Abstand zwischen a und b kleiner ist als der Abstand zwischen c und d. In der Sprache der Messbarkeitstheorie heißt das, man möchte ordinale Daten mithilfe einer ordinalen Skala repräsentieren, die metrische Eigenschaften erfüllt. Da im weiteren die Grundlegenden Skalen noch öfters vorkommen, definieren wir diese hier in der Einleitung kurz und intuitiv, ohne die genaue mathematische Definition zu nennen.

Definition 1 (Skalentypen)**.**
Die Skala ist eine Eigenschaft von Messdaten und enthält Information über deren Struktur. Dabei wird zwischen Vier Grundlegenden Skalen unterschieden, die sich jeweils mathematisch durch die zulässigen Operationen und die zugrundeliegende Struktur unterscheiden.

1. Nominalskala: Skala bei der man nur 1-zu-1 Zuordnungen macht. Die zulässigen Transformationen sind alle bijektiven Abbildungen. Beispiele sind z.B. Geschlechter (in Fragebögen z.B. 1 weibl. 2 männl. 3 divers) oder Postleizahlen.

2. Ordinalskala Zusätzlich zu 1-zu-1 Zuordnungen hat man eine Reihenfolge. Daher sind die zulässigen Transformationen die streng monotonen Abbildungen. Beispiele sind Noten (1 bis 6) oder Erdbebenstärke.

3. Intervallskala Zusätzlich zur Reihenfolge, spielen die Abstände zweier Werte eine Rolle. Damit sind die zulässigen Transformationen affin lineare Abbildungen $f(x) = ax + b, a > 0$. Beispiele sind die Celsius und Fahrenheit Skala.

4. Verhältnisskala Zusätzlich zu den Abständen in der Intervallskala, gibt es eine feste Null. Daher sind die zulässigen Transformationen alle linearen Abbildungen $f(x) = ax, a > 0$. Beispiele sind Länge oder die Kelvin Skala.

Da wir uns mit Ordinaldaten beschäftigen werden, spiele diese Skala und streng monotone steigende Abbildungen eine wichtige Rolle. Daher wird streng monoton steigend nur mit steigend bezeichnet.

In den Abschnitten 2 bis 4 geht es erstmal um die Mathematische Basis, die benötigt wird um die Hauptresultate aus Abschnitt 5 herzuleiten. Daher werden die wichtigsten Definitionen und Sätze in den Abschnitten 2 bis 4 erläutert aber nicht bewiesen. An ein paar Stellen sind die Feinheiten der Definitionen auch nicht so wichtig, dort wird dann auf den Anhang verwiesen, wo die genauen Definitionen von einigen Begriffen zu finden sind. Alle übersprungenen Beweise sind in [5] zu finden. Außer den Sätzen in [5] benötigen wir noch 2 Weitere Resultate die nicht Beweisen werden: [4] und [1], diese beiden Paper sind auf Moodle zu finden.

2. Proximity Structures und Metriken

In diesem Kapitel geht es erstmal darum die Mathematische Basis Aufzubauen. Wie aus der Einleitung hervorgeht, wollen wir jeweils zwei Paare von Stimuli auf ihre Ähnlichkeit vergleichen. Seien a, b, c und d solche Stimuli. Dann bedeutet die Notation $(a, b) \succsim (c, d)$, dass Stimuli a und b sich unähnlicher sind, als c und d. Analog bedeutet $\sim$, dass die Ähnlichkeit zweier Paare gleich ist.
Die Idee von einer proximity structure, ist die Mathematische Definition dieser Idee: eine Relation $\succsim$ auf einer Menge A die die Grundeigenschaften einer Metrik erfüllt.

Definition 2 (factorial proximity structure)**.**
Sei A eine Menge und $\succsim$ ist eine vierstellige Relation auf A, d.h. eine zweistellige Relation auf $A \times A$. $\langle A, \succsim \rangle$ ist eine proximity structure genau dann, wenn die folgenden Axiome für alle $a, b \in A$ gelten:

1. $\succsim$ ist eine schwache Ordnung von $A \times A$, d.h. sie ist zusammenhängend und transitiv.
2. $(a, b) \succsim (a, a)$ wenn $a \neq b$.
3. $(a, a) \sim (b, b)$ (Minimalität).
4. $(a, b) \sim (b, a)$ (Symmetrie).

Die Struktur ist factorial genau dann, wenn

$$A = \prod_{i=1}^{n} A_i.$$

Bemerkung 3. Zusammenhängend heißt, man kann alle distinkten Elemente vergleichen, also für eine Relation R auf einer Menge X gilt für $x, y \in X, x \neq y$ dann xRy oder yRx.

Da wir die Relation $\succsim$ durch eine Metrik repräsentieren wollen, lohnt es sich einen Begriff dafür einzuführen. Die Idee ist das, was man sich intuitiv vorstellt: man will dass die Metrik die Struktur von $\langle A, \succsim \rangle$ erhält, und die einzige Struktur die wir gegeben haben, ist die Ordnung jeweils zweier Paare.

Definition 4 (Darstellbarkeit von $\succsim$ durch Metrik)**.**
Sei A eine Menge und $\succsim$ eine vierstellige Relation auf A. Die Struktur $\langle A, \succsim \rangle$ ist genau dann durch eine Metrik darstellbar, wenn es eine reellwertige Funktion δ auf A gibt, so dass:

1. $\langle A, \delta \rangle$ ist ein metrischer Raum, d.h. δ ist eine Metrik auf A.
2. δ erhält die Relation $\succsim$, d.h. für alle $a, b, c, d \in A$

$$\delta(a, b) \geq \delta(c, d) \iff (a, b) \succsim (c, d).$$

3. Metriken mit additiven Segmenten

Da Metriken an sich wenig Struktur bieten, lohnt es sich noch eine Eigenschaft mehr mehr zu fordern.
Diese Eigenschaft ist segmental additivity und beschreibt, dass es zu zwei Punkten immer eine Kurve gibt, die diese verbindet und die Abstände additiv sind, das heißt auf dem Segment die Dreiecksungleichung mit Gleichheit gilt. Für uns reicht es zu wissen, dass wenn b auf dem Segment zwischen a und c liegt, dass dann $\delta(a,b)+\delta(b,c)=\delta(a,c)$ gilt. Für die genaue Definition siehe Anhang A. Am Ende dieses Abschnittes sind Beispiele für Metriken mit additiven Segmenten angegeben.

Da wir diese Additivität aus der proximity structure herleiten wollen, müssen wir Eigenschaften definieren, aus der man die Additivität herleiten kann. Dazu benötigen wir Folgenden Begriff:

Definition 5. Für $a,b,c \in A$ gilt die dreistellige Relation $(a,b,c\rangle$ wenn für alle $a',b',c' \in A$ gilt:

1. wenn $(a,b) \succsim (a',b')$ und $(a',c') \succsim (a,c)$, dann gilt $(b',c') \succsim (b,c)$
2. Wenn beide der Ungleichungen strikt sind, ist die resultierende Ungleichung auch strikt

Weiter gilt $\langle abc\rangle$ wenn $(a,b,c\rangle$ und $(c,b,a\rangle$ beide gelten.

Die Bedingungen, die notwendig sind, um eine Darstellung durch eine Metrik mit additiven Segmenten zu ermöglichen, sind in der folgenden Definition und dem Satz danach zusammengefasst. Der zweite Teil der Definition wird erst später, für das Hauptresultat (Satz 21) relevant, im Prinzip geht es darum dass man sich Grenzwerte und Cauchy Folgen anschauen kann, also der normale Begriff von Vollständigkeit.

Definition 6 (segmentally additive proximity structure)**.**
Eine proximity structure $\langle A, \succsim\rangle$, mit einer dreifachen Relation $\langle\rangle$, die in Bezug auf $\succsim$ definiert ist, ist segmentally additive genau dann, wenn die folgenden zwei Axiome für alle $a,c,e,f \in A$ gelten:

1. Wenn $(a,c) \succsim (e,f)$, dann gibt es ein $b \in A$, so dass $(a,b) \sim (e,f)$ und $\langle abc\rangle$. (segmental solvability)
2. Wenn $e \neq f$, dann gibt es $b_0,\ldots,b_n \in A$, so dass $b_0 = a, b_n = c$ und $(e,f) \succsim (b_{i-1},b_i), i = 1,\ldots,n$. (connectedness)

Eine segmentally additive proximity structure ist genau dann vollständig, wenn die folgenden zwei Axiome für alle $a,c,e,f \in A$ gelten:

3. Wenn $e \neq f$, dann gibt es $b,d \in A$, wobei $b \neq d$, so dass $(e,f) \succ (b,d)$. (Nichtdiskretheit der Abstände)

4. Sei $a_1, \dots, a_n, \dots$ eine Folge von Elementen von A, so dass für alle $e, f \in A$, wenn $e \neq f$, dann $(e,f) \succsim (a_i, a_j)$ für alle bis auf endlich viele Paare (i,j). Dann gibt es ein $b \in A$, so dass für alle $e, f \in A$, wenn $e \neq f$, dann $(e,f) \succsim (a_i, b)$ für alle bis auf endlich viele i. (Grenzwerte für Cauchy Folgen)

Der Folgende Satz liefert eine (bis auf positive Skalierung) eindeutige Metrik die die Relation repräsentiert und zusätzlich zur Ordnung, die Eigenschaft segmental solvability erhält.

Satz 7 (Repräsentationstheorem für segmentally additive proximity structures)**.**
Sei $\langle A, \succsim \rangle$ eine segmentally additive proximity structure. Dann ist $\langle A, \succsim \rangle$ durch eine Metrik darstellbar und zusätzlich gilt für alle $a, b, c \in A$ gilt:

1. $\langle abc \rangle \iff \delta(a,b) + \delta(b,c) = \delta(a,c)$.
2. *wenn δ' eine andere Metrik auf A ist, die die obigen Bedingungen erfüllt, dann gibt es ein $\alpha > 0$, so dass $\delta' = \alpha\delta$ (δ ist eine Verhältnisskala).*

3.1. Beispiele für Metriken mit additiven Segmenten

Beispiel 8 (Minkowski bzw. l^p bzw. power metric)**.**
Die l^p Metrik, die durch $\delta(a,b) = \left[\sum_{i=1}^n |\varphi_i(a_i) - \varphi_i(b_i)|^p\right]^{1/p}$ definiert ist, ist für $p \geq 1$ eine Metrik mit additiven Segmenten. Die Segmente sind genau die Verbindungslinien zwischen zwei Punkten, also alle Geraden.

Beweis. Seien $x, z \in \mathbb{R}^n$ und $0 \leq t \leq 1$. Dann gilt für $y = (1-t)x + tz$ (y ist genau dann auf dem Geradensegment von z zu x, wenn man y so darstellen kann):

$$\begin{aligned} d(x,y) + d(y,z) &= \left[\sum_{i=1}^n |x_i - (1-t)x_i - tz_i|^p\right]^{1/p} + \left[\sum_{i=1}^n |(1-t)x_i + tz_i - z_i|^p\right]^{1/p} \\ &= t\left[\sum_{i=1}^n |x_i - z_i|^p\right]^{1/p} + (1-t)\left[\sum_{i=1}^n |x_i - z_i|^p\right]^{1/p} \\ &= d(x,z). \end{aligned}$$

□

Bemerkung 9 (additive difference model)**.** Es fällt auf, dass die l^p Metrik eine besondere Form hat, die durch $\delta(x,y) = F^{-1}\left[\sum_{i=1}^n F(|x_i - y_i|)\right]$ gegeben ist. Man kann diese Form auf 4 Stufen abstrahieren:

1. Decomposability: Die Distanz ist eine Funktion von Beiträgen die Komponentenweise sind.

$$\delta(a,b) = F[\psi(a_1,b_1), \dots, \psi(a_n,b_n)] \tag{1}$$

2. Intradimensional subtractivity: Die Komponentenweise Beiträge sind Beträge der Differenzen in den Dimensionen.

$$\delta(a,b) = F[|\varphi_1(a_1) - \varphi_1(b_a)|, \ldots, |\varphi_n(a_n) - \varphi_n(b_n)|] \tag{2}$$

3. Interdimensionale additivity: Der Abstand ist eine Summe der Komponentenweise Beiträge.

$$\delta(a,b) = G\left[\sum_{i=1}^{n} \psi_i(a_i, b_i)\right] \tag{3}$$

4. additive difference model: Wenn alle 3 Eigenschaften gelten.

$$\delta(a,b) = G\left\{\sum_{i=1}^{n} F_i[|\varphi_i(a_i) - \varphi_i(b_i)|]\right\} \tag{4}$$

Falls die Elemente a, b aus der proximity structure sind, sind noch Folgende Eigenschaften von den Funktionen gefordert:

1. F ist steigend in allen Argumenten

2. ψ_i ist für alle $i = 1, ..., n$ reellwertig, symmetrischen und erfüllt $\psi_i(a_i, b_i) > \psi_i(a_i, a_i)$ wenn $a_i \neq b_i$

3. ϕ_i ist reellwertig für alle $i = 1, ..., n$

4. G und F_i sind steigend $(i = 1, ..., n)$

Man sieht, dass die l^p Metrik ein Spezialfall des additive difference models ist. Von besonderem Interesse sind Modelle wo die F_i alle über eine gemeinsame Funktion F (bis auf Streckung/Stauchung des Definitionsbereiches, $F_i(x) = F(t_i x), t_i > 0$) dargestellt werden können und $G = F^{-1}$ gilt. Diese Bedingungen erfüllt die l^p Metrik auch.
Hier kommt ein weiteres Beispiel für eine solche Metrik.

Beispiel 10 (p-exp Metrik Familie)**.**
Sei das F im additive difference model (4) gegeben durch

$$F(t) = q \cdot (e^{t \cdot r} - 1)^p \qquad q, r > 0, p \geq 1 \tag{5}$$

Dann ist das Modell eine Metrik.

Beweis. 1. Positivität: $\delta(a,b) = 0$ also $F(|a_i - b_i|) = 0$ für alle i, das ist nur der Fall wenn der Exponent Null ist, also $a_i = b_i$ gilt für alle i.
Anderes herum gilt, wenn $a \neq b$ ist, dass dann $a_i \neq b_i$ für mindestens ein i, woraus folgt dass $F_i(|a_i - b_i|) > 0$, also auch $\delta(a,b) > 0$.

2. Symmetrie: Symmetrie kommt daher, dass $|a_i - b_i| = |b_i - a_i|$ gilt.

3. Dreiecksungleichung: Die Dreiecksungleichung kommt aus dem Satz von Mullholland [4, S. 297]. Das Paper von Mulholland ist auf Moodle zu finden.

$\square$

Satz 11 (Mulholland)**.**
Die Ungleichung

$$f^{-1}\left[\sum_i f(|x_i + y_i|)\right] \leq f^{-1}\left[\sum_i f(|x_i|)\right] + f^{-1}\left[\sum_i f(|y_i|)\right].$$

gilt für alle $x, y \in \mathbb{R}^n, n = 1, 2, \dots$ *wann immer die folgenden drei Bedingungen erfüllt sind:*

1. f *ist stetig und streng monoton steigend, und* $f(0) = 0$.
2. f *ist konvex auf* $[0, \infty)$.
3. $\log f(e^x)$ *ist konvex für* $x \in (-\infty, \infty)$.

Beweis, dass 10 die Bedingungen von Mulholland erfüllt.
Zu Punkt 1: Stetigkeit und $f(0) = 0$ ist klar. Für Monotonie gilt, $e^{rt} - 1$ ist steigend, Potenzieren auch und multiplizieren mit $q > 0$ auch, damit ist die Komposition davon auch steigend. Zum Punkt 2: Selbes Argument wie bei Monotonie, nur mit Konvexität und Komposition von konvexen Funktionen ist wieder konvex.

Zum dritten Punkt, dass $\log F(e^t)$ konvex ist. Es gilt

$$\log q(e^{re^x} - 1)^p = \log q + p \cdot \log(e^{re^x} - 1)$$

Addition und Multiplikation mit $p \geq 1$ erhält Konvexität, also reicht es Konvexität von $\log(e^{re^x} - 1)$ zu zeigen.

Das zeigen wir über die zweite Ableitung, denn wenn die zweite Ableitung strikt positiv ist, ist die Funktion konvex.

Zweifaches Ableiten liefert

$$\frac{re^{re^x + x}\left(r(-e^x) + e^{re^x} - 1\right)}{\left(e^{re^x} - 1\right)^2} \overset{!}{>} 0$$

Wir sehen der Nenner ist immer positiv und ungleich 0, das heißt wir können den Nenner weg lassen. Genauso ist der Faktor $re^{re^x + x}$ immer positiv und kann somit auch weggelassen werden. Es bleibt also zu zeigen, dass

$$\begin{aligned} -re^x + e^{re^x} - 1 &> 0 \\ \iff e^{re^x} - re^x &> 1 \end{aligned}$$

gilt.

Dies kann man in 2 Schritten zeigen: 1. die Linke Seite ist streng monoton steigend und 2. der Grenzwert für $x \to -\infty$ ist 1. Somit gilt die Ungleichung für alle $x \in \mathbb{R}$.

1. Für die Ableitung gilt:

$$re^x \cdot e^{re^x} - re^x = re^x \cdot (e^{re^x} - 1) > 0$$

da $e^x > 0$ gilt.

2. $$\lim_{x \to -\infty} e^{re^x} - re^x = 1 - 0 = 1$$

Damit ist die Konvexität gezeigt, und die Behauptung folgt aus dem Satz von Mulholland.

□

Bemerkung 12 (l^p Metrik $\in$ p-exp Metrik Familie)**.**
Wenn wir in (5) für $q = r^{-p}$ einsetzen, erhalten wir im Grenzwert für $r \to 0$ die l^p Metrik. Das ist von Bedeutung, da dies mithilfe des Hauptresultates 21 zeigt, dass die Metriken der p-exp Familie keine Metriken mit additiven Segmenten sein können, obwohl man die l^p Metrik durch Grenzwertbildung aus dieser Familie darstellen kann. Das bedeutet, dass im Grenzwertprozess aus der Eigenschaft "additiv auf den Koordinatenachsen", die Eigenschaft "additiv auf beliebigen Liniensegmenten" wird.

Beweis. Setzt man $q = r^{-p}$ erhält man:

$$r^{-p}(e^{rt} - 1)^p = \left(\frac{e^{rt} - 1}{r}\right)^p$$

Für $r \to 0$ gehen sowohl Nenner, als auch Zähler, gegen 0 (wir können den Grenzwert in die Potenz reinziehen, da die Potenzfunktion stetig ist) und wir können den Satz von L'Hospital anwenden. Das liefert:

$$\lim_{r \to 0} \left(\frac{e^{rt} - 1}{r}\right)^p = \left(\lim_{r \to 0} \frac{e^{rt} - 1}{r}\right)^p = \left(\lim_{r \to 0} \frac{t \cdot e^{rt} - 1}{1}\right)^p = t^p$$

□

Nicht jede Funktion, die durch das additive difference model dargestellt werden kann, ist eine Metrik.

4. Additive difference proximity structures

In diesem Abschnitt geht es darum, wann eine factorial proximity structure durch das additive difference model (siehe 4) dargestellt werden kann. Die Axiome die dafür benötigt sind, werden hier nicht genauer dargestellt (siehe Anhang A), sonder nur genannt.

Definition 13 (additive difference structure)**.**
Eine factorial proximity structure $\langle A, \succsim \rangle$ mit $n \geq 3$ Faktoren ist eine additive difference structure, wenn die Folgenden Axiome erfüllt sind:

1. Betweenness
2. restricted solvability
3. Archimedean axiom
4. Independence

Diese Axiome liefern zusammen nach Folgendem Satz, dass die Struktur durch ein additive difference model darstellbar ist, welches die Ordnung behält. Die Axiome einzeln bzw. jeweils 2 oder 3 zusammen liefern die Bausteine des additive difference models, also decomposability, intradimensional subtractivity und interdimensional additivity. Falls $n = 2$ gilt, kann man eine Bedingung, die Thomsen condition, hinzufügen, um das selbe Resultat zu erhalten.

Satz 14 (Repräsentationstheorem für additive difference proximity structures)**.**
Sei $\langle A, \succsim \rangle$ eine additive-difference factorial proximity structure.
Dann gibt es reellwertige Funktionen $\varphi_1, \ldots, \varphi_n$, die auf $A_1, \ldots, A_n$ definiert sind, und steigende Funktionen $F_1, \ldots, F_n$, die jeweils auf $\mathbb{R}$ definiert sind, so dass für alle $a, b, c, d \in A$,

$$(a_1, \ldots, a_n, b_1, \ldots, b_n) \succsim (c_1, \ldots, c_n, d_1, \ldots, d_n)$$
$$\iff$$
$$\sum_{i=1}^{n} F_i(|\varphi_i(a_i) - \varphi_i(b_i)|) \geq \sum_{i=1}^{n} F_i(|\varphi_i(c_i) - \varphi_i(d_i)|).$$

Außerdem sind die φ_i Intervallskalen und die F_i Intervallskalen mit einer gemeinsamen Einheit ($i = 1, \ldots, n$).

5. Additive Difference Metrics

Satz 15 (Hinr. und notw. Bedingungen für additive difference Metrik)**.**
Sei $\langle A, \succsim \rangle$ additive-difference proximity structure sodass f die Darstellung der additive-difference Repräsentation nach Satz 14 ist, also

$$f(a, b) = \sum_{i=1}^{n} F_i(|\varphi_i(a) - \varphi_i(b)|).$$

Wir nehmen an, dass der Definitionsbereich jedes F_i ein reelles Intervall ist und $F_i(0) = 0$, $i = 1, \ldots, n$ (durch Transformation, weil F_i Intervallskalen sind).

Damit $\langle A, \succsim \rangle$ darstellbar durch eine Metrik ist und zusätzlich Additivität auf Koordinatenachsen gilt,

1. *ist es notwendig, dass es eine stetige, konvexe Funktion F gibt, so dass $F_i(x) = F(t_i x), t_i > 0, \forall x$ im Definitionsbereich von F_i, $i = 1, \ldots, n$; und*

2. *es ist ausreichend, wenn zusätzlich* $\log F(e^x)$ *konvex ist. (zusätzlich zu den Eigenschaften in 1.)*

Bemerkung 16 (Additiv auf den Koordinatenachsen)**.**
Additivität auf den Koordinatenachsen bedeutet hier: Wenn $a, b, c \in A$ sich nur auf einem Faktor A_i unterscheiden, dann gilt die $\triangle$-Ungleichung der Metrik mit Gleichheit. In Mathe Sprache heißt das:

$$a_j = b_j = c_j \forall j \neq i \text{ und } (a, c) \succsim (a, b), (b, c) \Rightarrow \delta(a, b) + \delta(b, c) = \delta(a, c)$$

Korollar 17 (Form von δ)**.**
Seien die Voraussetzungen wie in Satz 15. Dann wird die Metrik δ dargestellt durch

$$\delta(a, b) = F^{-1}\Big\{\sum_{i=1}^{n} F(|\varphi_i(a_i) - \varphi_i(b_i)|)\Big\} \tag{6}$$

wobei F stetig, konvex und steigend ist.

Beweis von 1.

Nach den Annahmen des Satzes existiert eine Metrik δ auf A und eine steigende Funktion G, so dass für alle $a, b \in A$,

$$\delta(a, b) = G\left(\sum_{i=1}^{n} F_i\left(|\varphi_i(a_i) - \varphi_i(b_i)|\right)\right).$$

Angenommen, a, b und c unterscheiden sich nur im i-ten Faktor und $a|b|c$ (Also wir können Additivität auf der i-ten Koordinatenachse für a, b, c benutzen). Daher gilt

$$G\left(F_i\left(|\varphi_i(a_i) - \varphi_i(b_i)|\right)\right) + G\left(F_i\left(|\varphi_i(b_i) - \varphi_i(c_i)|\right)\right) = G\left(F_i\left(|\varphi_i(a_i) - \varphi_i(c_i)|\right)\right).$$

Definiere $|\varphi_i(a_i) - \varphi_i(b_i)| = x$ und $|\varphi_i(b_i) - \varphi_i(c_i)| = y$, was ergibt

$$G[F_i(x)] + G[F_i(y)] = G[F_i(x + y)].$$

Definiere $H_i(x) = G[F_i(x)]$. Damit reduziert sich die obige Gleichung auf $H_i(x) + H_i(y) = H_i(x + y)$ für alle x, y im Definitionsbereich von $F_i, 1 \leq i \leq n$ und H_i muss linear sein. Da die Werte positiv sind, ist die einzige monotone Lösung dieser Gleichung $H_i(x) = t_i x$, $t_i > 0$. Da G steigend ist, existiert G^{-1} und es gilt

$$G^{-1}[H_i(x)] = G^{-1}[t_i x] = F_i(x)$$

wenn F_i definiert ist. Da G^{-1} unabhängig von i ist, können die F_i sich nur im Definitionsbereich bzw. im Wertebereich unterscheiden. ObdA können wir annehmen, dass der Wertebereich von F_n am größten ist. Da der Bereich von F_n den Bereich von F_i umfasst, können wir $F = F_n$ setzen und erhalten damit $F_i(x) = F_n(t_i x) = F(t_i x)$ für $t_i > 0$ und für jedes x im Definitionsbereich von F_i. Stetigkeit und Konvexität werden in den nächsten Lemmata bewiesen. $\square$

Lemma 18. Sei F wie in Satz 15. Dann gilt $F(\alpha+\beta+\gamma)+F(\gamma) \geq F(\alpha+\gamma)+F(\beta+\gamma)$.

Beweis. In der Dreiecksungleichung,

$$F^{-1}\left[\sum_{i=1}^{n} F(\alpha_i)\right] + F^{-1}\left[\sum_{i=1}^{n} F(\beta_i)\right] \geq F^{-1}\left[\sum_{i=1}^{n} F(\alpha_i+\beta_i)\right],$$

seien $\alpha_1 = \gamma, \alpha_2 = F^{-1}(F(\alpha+\gamma) - F(\gamma)), \alpha_j = 0$ für $3 \leq j \leq n$ und $\beta_1 = \beta, \beta_j = 0$ für $2 \leq j \leq n$. Damit gilt,

$$\begin{aligned} F(\alpha+\beta+\gamma) &= F\left[F^{-1}\left[\sum_{i=1}^{n} F(\alpha_i)\right] + F^{-1}\left[\sum_{i=1}^{n} F(\beta_i)\right]\right] \\ &\geq \sum_{i=1}^{n} F(\alpha_i+\beta_i) \quad \text{(durch die Dreiecksungleichung)} \\ &= F(\beta+\gamma) + F(\alpha+\gamma) - F(\gamma). \end{aligned}$$

□

Lemma 19. Sei F wie in Satz 15. Dann ist F stetig.

Beweis. Da F steigend ist, sind alle Unstetigkeiten Sprungstellen. Aus Analysis wissen wir, dass reelle monotone Funktionen höchstens abzählbar viele Sprungstellen (Unstetigkeitesstellen erster Art) haben. Dies führen wir auf einen Widerspruch.
Angenommen, $F(\alpha)$ ist die untere Grenze einer Sprungstelle, dann existiert $\epsilon > 0$, so dass für alle $\delta > 0$, $F(\alpha+\delta) - F(\alpha) > \epsilon$. Nach Lemma 4 gilt jedoch für jedes $\beta > 0$,

$$F(\alpha+\beta+\delta) - F(\alpha+\beta) \geq F(\alpha+\delta) - F(\alpha) > \epsilon,$$

und somit gibt es eine Lücke bei jedem $\alpha+\beta$, und F müsste überall unstetig sein. Dies ist unmöglich, also muss F stetig sein. □

Lemma 20.
Sei F eine streng monoton steigende Funktion, die auf den nichtnegativen reellen Zahlen definiert ist. Dann ist F genau dann konvex, wenn für alle $\alpha, \beta, \gamma \geq 0$ die Ungleichung aus 18 gilt, also

$$F(\alpha+\beta+\gamma) + F(\gamma) \geq F(\alpha+\gamma) + F(\beta+\gamma).$$

Beweis. $\Rightarrow$: Angenommen, F ist konvex, und sei $p = \alpha/(\alpha+\beta)$, dann gilt

$$\begin{aligned} p(\alpha+\beta+\gamma) + (1-p)\gamma &= \alpha+\gamma \\ (1-p)(\alpha+\beta+\gamma) + p\gamma &= \beta+\gamma. \end{aligned}$$

Damit gilt wegen Konvexität

$$
\begin{aligned}
F(\alpha+\gamma)+F(\beta+\gamma) &= F\left[p(\alpha+\beta+\gamma)+(1-p)\gamma\right]+F\left[(1-p)(\alpha+\beta+\gamma)+p\gamma\right] \\
&\leq pF(\alpha+\beta+\gamma)+(1-p)F(\gamma)+(1-p)F(\alpha+\beta+\gamma)+pF(\gamma) \\
&= F(\alpha+\beta+\gamma)+F(\gamma).
\end{aligned}
$$

$\Leftarrow$: Um die Umkehrung zu beweisen, benutzen wir die Stetigkeit von F aus Lemma 19. Angenommen, $\alpha > \beta$, dann

$$
\begin{aligned}
F(\alpha)+F(\beta) &= F\left(\frac{\alpha-\beta}{2}+\frac{\alpha-\beta}{2}+\beta\right)+F(\beta) \\
&\geq 2F\left(\frac{\alpha-\beta}{2}+\beta\right) \quad \text{(nach Lemma 4)} \\
&= 2F\left(\frac{\alpha+\beta}{2}\right),
\end{aligned}
$$

was äquivalent zur Konvexität für ein stetiges F ist. □

Satz 21 (Eindeutigkeitssatz zu l^p Metriken)**.** *Sei $\langle A, \succsim\rangle$ eine vollständige, segmentally additive proximity structure, die auch eine additive difference structure ist.*
Ferner sei δ homogen, also $\delta(x, x+t(z-x)) = t\delta(x,z)$, für $t>0$.
Dann gibt es ein eindeutiges $r \geq 1$ und reellwertige Funktionen φ_i, die auf A definiert sind, $i = 1,\ldots,n$, die bis auf eine positive lineare Transformation eindeutig sind, so dass für alle $a,b,c,d \in A$,

$$(a,b) \succsim (c,d) \iff \delta(a,b) \geq \delta(c,d),$$

wobei

$$\delta(a,b) = \left[\sum_{i=1}^{n} |\varphi_i(a)-\varphi_i(b)|^r\right]^{1/r}.$$

Bemerkung 22. Dieser Satz ist ursprünglich ohne die Zusatzannahme der Homogenität von δ formuliert. Jedoch konnten wir den Beweis an einer Stelle nicht nachvollziehen bzw. denken dass er möglicherweise sogar falsch ist, und nehmen daher zusätzlich diese Eigenschaft an.

Beweis. Der Beweis passiert in mehreren Schritten. Zuerst sagen wir etwas über die Definitions- und Zielbereiche der Funktionen F_i und finden damit Grenzen für den Definitionsbereich von F. Danach argumentieren wir, dass wir F bis auf die positiven reellen Zahlen $[0,\infty)$ erweitern können und dass diese Erweiterung eine Funktionsgleichung impliziert, die nur von $f(x) = kx^p$ erfüllt wird. Im letzten Schritt zeigen wir, dass F auf ganz $\mathbb{R}^+$ gleich der Erweiterung sein muss und damit δ eine l^p Metrik sein muss.

Nach Satz 15 liefert das additive difference Model

$$\delta(a,b) = G\left(\sum_{i=1}^{n} F_i\left(|\varphi_i(a_i)-\varphi_i(b_i)|\right)\right)$$

nur dann eine Metrik, wenn $F_i(\alpha) = F(t_i\alpha)$ und $F = G^{-1}$ gilt. Daher ist F^{-1} für alle Zahlen $\sum_{i=1}^{n} F_i\left(|\varphi_i(a_i) - \varphi_i(b_i)|\right)$ definiert, und jedes F_i stimmt mit F auf seinen Definitionsbereich überein (bis auf den Faktor t_i, der das Argument multipliziert und den Definitionsbereich anpasst).
Da die Definitionsbereiche von F_i Intervalle sind (weil δ eine Metrik mit additiven Segmenten ist), können wir $\omega > 0$ wählen, so dass für jedes $i = 1, \dots, n$ und jedes α mit $0 \le \alpha \le \omega/t_i$ gilt, dass α im Bereich von F_i liegt.
Insbesondere, wenn wir $\omega_i = \omega/t_i$ setzen, ist $F^{-1}\left[\sum_{i=1}^{n} F_i(\omega_i)\right] = F^{-1}\left[nF(\omega)\right] =: \Omega$ definiert. Damit ist $F = G^{-1}$ mindestens im Intervall $[0, \Omega]$ definiert. Wir können F nun wie folgt erweitern.

Für $\alpha < \omega$ gilt mit der Homogenität

$$\begin{aligned} F^{-1}[nF(\alpha)] &= F^{-1}[nF(\alpha\omega/\omega)] \\ &= (\alpha/\omega)F^{-1}[nF(\omega)] \\ &= \alpha\Omega/\omega. \end{aligned}$$

Das liefert durch Einsetzen der korrekten Werte für α ($\tilde{\alpha} = \alpha\frac{\omega}{\Omega}$ und $\tilde{\alpha} = F^{-1}(\frac{\alpha}{n})$)

$$F(a) = nF(a\omega/\Omega), \quad 0 \le a \le \Omega; \tag{7}$$

$$F^{-1}(a) = (\Omega/\omega)F^{-1}(a/n), \quad 0 \le a \le nF(\omega). \tag{8}$$

Diese beiden Gleichungen können wir jetzt nutzen, um den Definitionsbereich von F und F^{-1} zu erweitern. Wir nutzen (7) indem wir F durch die rechte Seite definieren (für Werte, bei denen die rechte Seite definiert ist, also für $0 \le a \le \Omega^2/\omega$). Da $\Omega^2/\omega > \Omega$ gilt erweitert dies F wirklich über das Intervall $[0, \Omega]$ hinaus.
Entsprechend erfüllt F^{-1} Gleichung (8) für das erweiterte Intervall $0 \le a \le nF(\Omega) = n^2F(\omega)$.
Darüber hinaus gilt die Homogenität nun für $0 \le t\alpha_i \le \Omega$:

$$
\begin{aligned}
F^{-1}\left[\sum_{i=1}^{n} F(t\alpha_i)\right] &= F^{-1}\left[\sum_{i=1}^{n} nF(t\alpha_i\omega\backslash\Omega)\right] \quad \text{(nach Gleichung 7)}\\
&= F^{-1}\left[n\sum_{i=1}^{n} F(t\alpha_i\omega\backslash\Omega)\right]\\
&= \frac{\Omega}{\omega}F^{-1}\left[\sum_{i=1}^{n} F(t\alpha_i\omega\backslash\Omega)\right] \quad \text{(nach Gleichung 8)}\\
&= \frac{\Omega}{\omega}tF^{-1}\left[\sum_{i=1}^{n} F(\alpha_i\omega\backslash\Omega)\right] \quad \text{(wegen Homogenität, da } \alpha_i\omega\backslash\Omega \leq \omega\text{)}\\
&= tF^{-1}\left[n\sum_{i=1}^{n} F(\alpha_i\omega\backslash\Omega)\right] \quad \text{(nach Gleichung 8)}\\
&= tF^{-1}\left[\sum_{i=1}^{n} nF(\alpha_i\omega\backslash\Omega)\right]\\
&= tF^{-1}\left[\sum_{i=1}^{n} F(\alpha_i)\right] \quad \text{(nach Gleichung 7)}
\end{aligned}
$$

Durch wiederholte Anwendung erweitern sich die Intervalle, in denen F definiert ist und Homogenität gültig ist, jedes Mal um einen Faktor Ω/ω. Somit können wir F auf $[0,\infty)$ erweitern sodass die Homogenität erhalten bleibt.

Die Funktion, die durch das erweitern entstanden ist, muss somit mindestens auf dem Intervall $[o,\infty)$ mit der ursprünglichen Funktion F übereinstimmen. Außerdem gilt mit Gleichung 8 durch Einsetzen der Mittelwertsfunktion $F^{-1}\left[\left(\frac{1}{n}\sum_{i=1}^{n} F(\alpha_i)\right)\right]$:

$$
F^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} F(t\alpha_i)\right] = tF^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} F(\alpha_i)\right].
$$

Nach [3, Satz 84] sind die einzigen monotonen Funktionen, die diese Gleichung für $0 \leq \alpha < \infty$ erfüllen, Funktionen der Form $k\alpha^r$ (siehe auch [1, S. 153]).

Damit gilt, dass $F(\alpha) = k\alpha^r$ mindestens auf dem Intervall $[0,\Omega]$ gelten muss. Aber jetzt können wir, für jedes $|\varphi_i(a_i) - \varphi_i(b_i)|$, ein $t > 0$ so klein finden, dass $t|\varphi_i(a_i) - \varphi_i(b_i)| < \omega, i = 1,\dots,n$. Anwenden von der Monotonie liefert (wir können annehmen, dass $t \leq 1$)

$$
\begin{aligned}
\delta(a,b) &= F^{-1}\left(\sum_{i=1}^{n} F\left(|\varphi_i(a_i) - \varphi_i(b_i)|\right)\right) \\
&= \frac{1}{t} F^{-1}\left(\sum_{i=1}^{n} F\left(t|\varphi_i(a_i) - \varphi_i(b_i)|\right)\right) \\
&= \frac{1}{t}\left(\sum_{i=1}^{n} \left(t|\varphi_i(a_i) - \varphi_i(b_i)|\right)^r\right)^{1/r} \qquad \text{weil } t|\varphi_i(\alpha_i) - \varphi_i(\beta_i)| \leq \omega \\
&= \left(\sum_{i=1}^{n} |\varphi_i(a_i) - \varphi_i(b_i)|^r\right)^{1/r}
\end{aligned}
$$

Damit ist δ eine l^p Metrik, und durch Satz 14 ist $F = k \cdot \alpha^r$ und damit $F_i = k(t_i\alpha)^r$ oder $F_i(\alpha) = k\alpha^r$, wenn wir die t_i in die φ_i einbauen. Schließlich folgt die Einschränkung $1 \leq r < \infty$ aus der Dreiecksungleichung oder, genauer gesagt, aus der Minkowski-Ungleichung.

□

Literatur

[1] J. Aczél, *Lectures on functional equations and their applications*, Academic press, 1966.

[2] I. Breinbauer, *Bernhard orth: Einführung in die theorie des messens. kohlhammer verlag, stuttgart/berlin/köln/mainz 1974, 131 seiten, 15, – dm (kohlhammer standards psychologie, basisbücher und studientexte).*, Vierteljahrsschrift für wissenschaftliche Pädagogik, 52 (1976), pp. 106 – 107.

[3] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, 1952.

[4] H. P. Mulholland, *On Generalizations of Minkowski's Inequality in the Form of a Triangle Inequality*, Proceedings of the London Mathematical Society, s2-51 (1949), pp. 294–307.

[5] P. Suppes, *Foundations of Measurement: Volume 2*, vol. 2, Elsevier, 2014.

[6] A. Tversky and D. H. Krantz, *The dimensional representation and the metric structure of similarity data*, Journal of mathematical psychology, 7 (1970), pp. 572–596.

A. Anhang

A.1. Metrik mit additiven Segmenten

Wie im Abschnitt oben, ist die Idee dass man zwei beliebige Punkte mit einer Kurve verbinden kann, auf der Distanzen additiv sind. Dies wird wie folgt mathematisiert:

Definition 23. Segment Sei (X, δ) ist ein metrischer Raum. Eine parametrische Kurve f in X ist eine stetige Funktion $\mathbb{R} \supseteq [\alpha, \beta] \to X$. Die parametrische Kurve f verbindet ihren Anfangspunkt $f(\alpha)$ und ihren Endpunkt $f(\beta) = \beta$.

1. Sei f eine parametrische Kurve in (X, δ) mit Definitionsbereich $[\alpha, \beta]$. Die Länge $L(f)$ von f ist definiert als

$$L(f) = \sup \left\{ \sum_{i=1}^{n} \delta\left(f(\alpha_{i-1}), f(\alpha_i)\right) \right\}$$

   über alle endlichen Sequenzen $\alpha_0, \alpha_1, \ldots, \alpha_n$ mit $\alpha = \alpha_0 \leq \alpha_1 \leq \ldots \leq \alpha_n = \beta$.

2. Wenn $L(f) < \infty$, dann ist f rektifizierbar.

3. Eine rektifizierbare Kurve ist ein segment, wenn die Länge gleich der Distanz von Start- und Endpunkt ist.

Definition 24. segmentally additive Metric space Ein metrischer Raum (M, δ) heißt segmentally additive, wenn zwei Punkte immer durch ein Segment verbunden werden können.

A.2. Additive Difference Proximity Structure Axioms

Definition 25 (betweenness axiom)**.** Sei $\langle A, \succsim \rangle$ eine factorial proximity structure. Sei $\succeq_i$ die induzierte Ordnung auf A^2. Wir sagen, dass b zwischen a und c liegt, bezeichnet als $a|b|c$, genau dann, wenn

$$(a_i, c_i) \succsim_i (a_i, b_i) \quad \text{für jedes } i.$$

Wir sagen, dass $\langle A, \succsim \rangle$ die betweenness axioms erfüllt, genau dann, wenn die folgenden Bedingungen für alle $a, b, c, d, a', b', c' \in A$ gelten:

1. Angenommen a, b, c, d unterscheiden sich höchstens auf einem Faktor, und $b \neq c$, dann
   (i) wenn $a|b|c$ und $b|c|d$, dann $a|c|d$ und $a|b|d$,
   (ii) wenn $a|b|c$ und $b|c|d$, dann $a|b|d$ und $b|c|d$

2. Angenommen a, b, c, a', b', c' unterscheiden sich höchstens auf einem Faktor, $a|b|c, \quad a'|b'|c'$ und $(b, c) \sim (b', c')$. Dann

$$(a, c) \succsim (a', c') \iff (a, b) \succsim (a', b').$$

Definition 26 (Restricted solvability)**.** Eine factorial proximity structure $\langle A, \succsim \rangle$ erfüllt restricted solvability, falls für alle $e, f, d, a \in A$ gilt:
wenn $(d, a) \succsim (e, f) \succsim (d, c)$ dann existiert ein $b \in A$, so dass $a|b|c$ und $(e, f) \succsim (d, b)$.

Definition 27 (Arcimedean axtiom)**.** Eine proximity structure $\langle A, \succsim \rangle$ erfüllt das Archimedean Axiom, falls für alle $a, b, c, d \in A$ mit $a \neq b$ für jede Folge $e^{(k)} \in A, k = 0, 1, \ldots$ mit $e^{(0)} = c$ gilt, dass $(e^k, e^{(k+1)}) \succ (a, b)$ und $(c, d) \succ (c, e^{(k+1)} \succ (c, e^{(k)})$ für alle k endlich ist.

Definition 28 (independence)**.** Eine factorial proximity structure $\langle A, \succsim \rangle$ erfüllt independence, falls Folgendes für alle $a, b, c, d, a', b', c', d' \in A$ gilt:

1. für alle der Paare (a, a'), (b, b'), (c, c'), (d, d') jeweils zwei Elemente identisch auf einem Faktor sind und
2. für alle der Paare (a, c), (b, d), (a', c'), (b', d') jeweils zwei Elemente identische Komponenten in den restlichen Faktoren haben, dann

$$(c, d) \succsim (c', d') \iff (a, b) \succsim (a', b').$$